

Improving sample efficiency of high dimensional Bayesian optimization with MCMC

Zeji Yi^{1*}

Yunyue Wei^{1*}

Chu Xin Cheng^{2*}

Kaibo He¹

Yanan Sui¹

¹*Tsinghua University*, ²*California Institute of Technology*.

ZEJIY@ANDREW.CMU.EDU

WEIYY20@MAILS.TSINGHUA.EDU.CN

CCHENG2@CALTECH.EDU

HKB21@MAILS.TSINGHUA.EDU.CN

YSUI@TSINGHUA.EDU.CN

Abstract

Sequential optimization methods are often confronted with the curse of dimensionality in high-dimensional spaces. Current approaches under the Gaussian process framework are still burdened by the computational complexity of tracking Gaussian process posteriors and need to partition the optimization problem into small regions to ensure exploration or assume an underlying low-dimensional structure. With the idea of transiting the candidate points towards more promising positions, we propose a new method based on Markov Chain Monte Carlo to efficiently sample from an approximated posterior. We provide theoretical guarantees of its convergence in the Gaussian process Thompson sampling setting. We also show experimentally that both the Metropolis-Hastings and the Langevin Dynamics version of our algorithm outperform state-of-the-art methods in high-dimensional sequential optimization and reinforcement learning benchmarks.

Keywords: Bayesian Optimization, Markov Chain Monte Carlo, High Dimensional Optimization

1. Introduction

With broad applications in real-world engineering problems, black-box function optimization is an essential task in machine learning (Snoek et al., 2012; Hernández-Lobato et al., 2017). These non-convex sequential optimization problems often lack gradient information. Bayesian optimization (BO) is a popular sampling-based online optimization approach for solving expensive optimization problems. It has been successfully applied to problems such as online learning and sequential decision-making. BO builds a surrogate model for modeling the objective function and optimizes the acquisition function to propose new samples.

Similar to many other numerical problems, BO algorithms are also susceptible to the curse of dimensionality. The search space would grow exponentially large as the function dimension increases and become intractable under a limited computation budget. Common acquisition functions also tend to over-explore the uncertainty boundary region and lack exploitation in high dimensional input space (Oh et al., 2018). Recent developments in high dimensional BO include constructing trust regions and space partitions to improve the probability of sampling in promising regions, which effectively ameliorates the problem of over-exploration (Eriksson et al., 2019; Wang et al., 2020). To evaluate the acquisition function on a continuous domain, these methods often discretize

*These authors contributed equally to this work. Code and Appendices are available at <https://drive.google.com/drive/folders/1fLUHIduB3-pr78Y1YOhhNtsDegaOqLNU?usp=sharing>.

the search space using Sobolev sequence (Sobol', 1967). However, this kind of discretization could be inadequate in high dimensional space as the size of the discretization set is limited, hindering the exploitation of potential good regions.

In our setting, we investigate probabilistic acquisition such as Thompson sampling (TS) for optimizing unknown function modeled with Gaussian processes (GP). Yet, directly applying algorithms that perform well in low-dimensional problems to high-dimensional domains often suffers due to the high-computation demand and the phenomenon of over-exploration. Our proposed method aims to improve optimization performance in high-dimensional spaces by adapting to the search domain and specifically discretizing it in promising regions. To achieve this, we utilize Markov Chain Monte Carlo (MCMC) for BO, which is a widely-used technique known for its ability to effectively sample from high-dimensional posterior distributions. We introduce MCMC-BO, a method that transitions candidate points from their original positions towards the approximated stationary distribution of TS. By only tracking a batch size number of points during transitions, MCMC-BO significantly reduces the storage requirements in comparison to an entire mesh while still maintaining theoretical performance guarantees. It can be easily generalized to different scenarios with few additional hyper-parameters. Based on the basic MH version of MCMC-BO. We also propose the Langevin dynamics (LD) version which transits following the gradient of the log-likelihood. Overall, MCMC-BO can serve as a versatile algorithm that can be linked to most existing BO methods providing a posterior for any candidate points for continuous problems. We summarize our contributions as follows:

- We propose MCMC-BO, a Bayesian optimization algorithm which performs adaptive local optimization in high-dimensional problems that achieves both time and space efficiency.
- We provide theoretical guarantees on the convergence of MCMC-BO. To our knowledge, this is the first regret bound on high-dimensional Bayesian optimization problem which can deal with the scaling of dimensions with limited candidate points per round and avoid the overuse of memory.
- We experimentally show that MCMC-BO, combined with both MH and LD, outperforms other strong baselines on various high-dimensional tasks.

2. Related Work

The optimization of black-box functions has been broadly used in many scenarios, such as hyper-parameter tuning and experimental design (Snoek et al., 2012; Hernández-Lobato et al., 2017). Evolutionary algorithms are a class of methods with a long history and good performance to tackle the black-box optimization problem. One representative is CMA-ES, which adaptively adjusts the

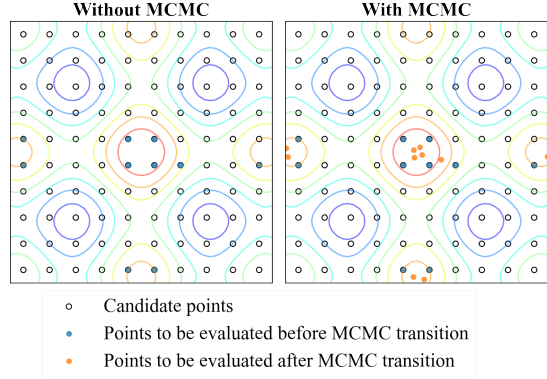


Figure 1: Illustration of MCMC-BO. The contours are 2d Rastrigin function. **(Left)**: BO algorithms propose points to be sampled. The optimization performance is restricted by insufficient discretization. **(Right)**: Points are adjusted by MCMC-BO, reaching regions with higher value.

covariance matrix to generate new samples (Hansen et al., 2003). These kinds of problems can also be formalized under the Bayesian optimization framework (Shahriari et al., 2015; Frazier, 2018), which is also more efficient. However, conventional BO algorithms are limited to low-dimensional cases and a small number of observations. There are two major components in BO: modeling and performing acquisition. To improve scalability and efficiency, various methods have been proposed for these two aspects. Regarding modeling, there exist a large amount of works which scale BO to higher dimensions via approximated GP (Hensman et al., 2013), dimension reduction (Nayebi et al., 2019) or function decomposition (Wang et al., 2018). We leave the detailed discussion of this part in the appendix.

In the acquisition phase, balancing the exploration-exploitation trade-off is important. High-dimensional optimization problems often suffer from inadequate exploitation. To prevent over-exploring uncertain points near the domain boundary, methods like strong priors, dimension dropout, and cylindrical kernels have been proposed to allocate more sample points in the center region (Eriksson and Jankowiak, 2021; Li et al., 2018; Oh et al., 2018). Full exploration in high-dimensional space requires unrealistic computational resources and time. Therefore, adequate exploitation of optimal points determines final performance. Previous algorithms for optimizing high-dimensional spaces include TuRBO and LA-MCTS. TuRBO limits the candidate set to a small box around the best points and adjusts the box based on later evaluations (Eriksson et al., 2019). LA-MCTS partitions the region with a constructed tree, recommending promising regions to improve algorithm performance (Wang et al., 2020). Both methods focus on dividing promising regions, but maintaining a high-precision discretized grid for such regions is costly and can limit computation storage. The trade-off between regret and storage emphasizes the need for improvements in current methods. Therefore, we propose attention to the ability to congregate candidate points around the optimum in our algorithm.

Focusing on probabilistic acquisition, which involves using stochastic samples from the model posterior, sampling from the posterior distribution can be a bottleneck, especially for problems without a closed form expression for its posterior distribution. To address this, approximate sampling techniques are commonly used to generate samples from posterior approximations (Chapelle and Li, 2011; Lu and Van Roy, 2017). These approaches have shown effective practical performance (Riquelme et al., 2018; Urteaga and Wiggins, 2017). However, maintaining performance over arbitrary time horizons while using approximate sampling remains unclear. Recent works have explored the application of LD, a stochastic sampling approach, in Multi-armed Bandit (MAB) problems (Mazumdar et al., 2020; Xu et al., 2022). Instead of restricting to a parameterized utility function and optimizing the parameter space, we directly employ LD on the continuous and infinite action space. We believe this approach is more general and effective as it allows direct acquisition of candidate points without the need for optimization in the high-dimensional space.

3. Background

3.1. Modeling with Gaussian processes

Black-box optimization under uncertainty aims to find

$$\mathbf{x}^* \in D \text{ such that } f(\mathbf{x}^*) \geq f(\mathbf{x}), \quad \forall \mathbf{x} \in D \quad (1)$$

where $f : D \rightarrow \mathbb{R}$ is an unknown function. In BO, the unknown objective f is viewed as a probability distribution, and GP emerges as an effective way to estimate the unknown function, maintain

an uncertainty estimate, and can be sequentially updated with available information. A GP is characterized by its mean function $\mu(\cdot)$ and kernel $k(\cdot, \cdot)$. The prior distribution of $f(x)$ is assumed to be $\mathcal{N}(0, k(x, x))$. Under this prior, the sample points $A_T := [x_1, \dots, x_T]$ and the observations $[y_1, \dots, y_T]$ follows the multivariate distribution $\mathcal{N}(0, K_T + \sigma^2 I)$, where $K_T = [k(x, x')]_{x, x' \in A_T}$. Denote $k_T(x) = [k(x_i, x)]_{i=1:T}^T \in \mathbb{R}^{T \times 1}$, the posterior distribution over f is thus Gaussian with mean $\mu_T(x)$ and covariance $k_T(x, x')$ that satisfy: $\mu_T(x) = k_T(x)^T (K_T + \sigma^2 I)^{-1} y_T$, $k_T(x, x') = k(x, x') - k_T(x)^T (K_T + \sigma^2 I)^{-1} k_T(x')$

3.2. Gaussian process and Thompson sampling

In online learning tasks where the actions x_t have to be sequentially chosen at each time step t , a common metric for measuring the performance is regret, which is defined using the optimal action $x^* := \arg \max_{x \in D} f(x)$ such that $r_t := f(x^*) - f(x_t)$ for the instantaneous regret at time t . The goal of optimization is to minimize the sum of r_t , which entails finding optimal points while effectively exploring the sample space. To achieve this, we build upon ideas inspired by TS, which is to select the best point using a stochastic sample from the most recent model. TS is a classical MAB algorithm that can balance the exploitation-exploration trade-off. In every round, the algorithm estimates the reward of each arm by sampling from their respective distributions and plays the arm with the highest estimate. Thus, exploitation of optimal arms is achieved by heuristically playing the optimal arm, while exploration is guaranteed by the intrinsic uncertainty of probabilistic sampling. With an evolving GP, the algorithm first samples $f_t \sim GP(\mu_t(x), k_t(x, x'))$ and selects $x_t := \arg \max_{x \in \Omega} f_t(x)$ as the candidate points. Chowdhury and Gopalan gave a theoretical analysis for the algorithm GP-TS, which combines GP and TS to optimize stochastic MAB problems ([Chowdhury and Gopalan, 2017](#)).

3.3. Markov Chain Monte Carlo

Markov Chain Monte Carlo is used in GP modeling to approximate intractable probability distributions ([Hensman et al., 2015](#)). Similarly, we hope to utilize MCMC to sample from a distribution on the action set corresponding to our acquisition approach.

Metropolis-Hastings (MH) is an MCMC sampling algorithm that can sample from a target distribution $\pi(x)$, $x \in \mathcal{X}$ known up to a constant, i.e. when we have knowledge of $\pi_d(x) = c \cdot \pi(x)$ ([Metropolis et al., 1953](#)). With a given proposal distribution $q(\cdot, x)$, the algorithm samples a candidate point $y \in \mathcal{X}$ given the current value x according to $q(y, x)$. Then the Markov Chain transitions to y with acceptance probability $\alpha(x, y) = \min \left\{ 1, \frac{\pi_d(y)q(y, x)}{\pi_d(x)q(x, y)} \right\}$. This ratio yields a high acceptance probability when the proposed point is likely to be a better choice than the current point, so the points transit towards regions with higher function values. The MH algorithm is guaranteed to converge to a stationary distribution that is exactly $\pi(\cdot)$.

Langevin dynamics (LD), known to converge to the steady state distribution $\pi(x)$, provides another efficient iterative sampling procedure. From any arbitrary point x_0 sampled from a prior distribution, the update rule is given by $x_{i+1} = x_i + \epsilon \cdot \nabla \log \pi(x) + \sqrt{2\epsilon} z_i$, $z_i \sim \mathcal{N}(0, I)$.

Directly applying these two MCMC algorithms is challenging, as we do not have access to the complicated target distribution $\pi(x)$ induced by our probabilistic acquisition approach. In Section 4, we propose estimates for the acceptance probability $\alpha(x, y)$ and $\nabla \log \pi(x)$ that allows us to sample from the posterior distribution.

4. Algorithm Design

We propose MCMC-BO, a novel algorithmic framework that combines MCMC methods with GP to evaluate more points effectively while maintaining a dense, time-varying discretization near promising regions for optimization.

The key idea of our algorithm is inspired by TS in which points with larger probability of being optimal are chosen next. However, probabilistic acquisition in this manner is computationally demanding due to the required size of the discretization set and the complexity of the induced distribution. Coarse partitioning of the input domain naturally leads to regret, as the best possible point in the mesh may still be far from the optimum. Consequently, a large number of points is necessary to ensure a fine partition. In high-dimensional spaces, this need for numerous points can result in large matrices for sampling in stochastic acquisition approaches and slow predictions. This presents a dilemma: a coarse mesh leads to significant regret, while a fine mesh makes sampling unrealistic. To efficiently sample from an intractable distribution and identify better candidate points, we employ MCMC.

MCMC-BO offers great flexibility in choosing different MCMC algorithms to sample from the GP posterior. We provide two MCMC subroutines using MH and LD. Our proposed method is compatible with any existing BO algorithms with probabilistic acquisition functions over a continuous domain (see Appendix). The MH algorithm introduced in Section 3.3 can serve as a overarching framework using various acceptance ratios, and we propose an acceptance ratio that approximates the transition probability induced by the acquisition approach.

Algorithm 1: MCMC-BO

Input: Initial dataset \mathcal{D}_0 , batch size m , MCMC transition number N , MCMC transition parameter Θ

```

1 for  $t = 0, 1, \dots$  do
2   Update posterior distribution  $\text{GP}(\mu_t(x), k_t(x, x'))$  on  $f$  using  $\mathcal{D}_t$ 
3   Create discretized candidate points  $\mathbf{x}_t^0$  from continuous search domain
   // MCMC transition start
4   for  $i = 0$  to  $N - 1$  do
5     for  $k = 0$  to  $m - 1$  do
6       | Perform [MCMC routine] using  $\text{GP}(\mu_t(x), k_t(x, x'))$  on  $\mathbf{x}_{tk}^i$  with parameter  $\Theta$ 
7     end
8   end
   // MCMC transition end
9   Observe  $\mathbf{y}_t = f(\mathbf{x}_t^N)$ 
10   $\mathcal{D}_{t+1} \leftarrow \mathcal{D}_t \cup (\mathbf{x}_t^N, \mathbf{y}_t)$ 
11 end
```

TS implicitly induces a distribution on each candidate point $x \in D_t$, with the probability of being selected as $\mathbb{P}(f_t(x) \geq f_t(x'), \forall x' \in D_t)$. Let x_p, x_o be two candidate points after n samples denoted by history $\mathcal{H} = \{(x_t, y_t)\}_{t=1}^n$. We can sample according to the distribution of TS using MH with an acceptance rate $\alpha = \min \left\{ 1, \frac{\mathbb{P}(f_t(x_p) \geq f_t(x), \forall x \in D_t) \cdot q(x_p | x_o)}{\mathbb{P}(f_t(x_o) \geq f_t(x), \forall x \in D_t) \cdot q(x_o | x_p)} \right\}$, where, D_t represents the discretization set of the domain D at time t , and x_p is determined by the transition kernel $q(\cdot | \cdot)$. However, computing the probability involved in the ratio, $\mathbb{P}(f_t(x_p) \geq f_t(x), \forall x \in D_t)$, requires

marginalizing over the joint distribution of a high-dimensional multivariate normal distribution, which is intractable. Therefore, we propose an alternative approach that captures a similar idea. We define the acceptance probability as $\alpha = \min \left\{ 1, \frac{\mathbb{P}(f_t(x_p) \geq f_t(x_o))q(x_p|x_o)}{\mathbb{P}(f_t(x_o) \geq f_t(x_p))q(x_o|x_p)} \right\}$. From GP regression, we can view the surrogate function values $y_p = f_t(x_p)$, $y_o = f_t(x_o)$ associated to selected points $\mathbf{x} := (x_p, x_o)$ as a Gaussian random vector with joint distribution as specified in Section 3.1. Using linear transformations of a Gaussian random vector, let $c^T = [1 \quad -1]$, $K_n = [k(x_{1:n}, x_{1:n})] \in \mathbb{R}^{n \times n}$, and $k_n(\mathbf{x}) = [k(x_{1:n}, x_p), k(x_{1:n}, x_o)] \in \mathbb{R}^{n \times 2}$, we have that $y_p - y_o \sim \mathcal{N}(c^T \mu, c^T \Sigma c)$, where $c^T \mu = (k(x_p, x_{1:n}) - k(x_o, x_{1:n}))^T (K_n + \sigma^2 I)^{-1} y_{1:n}$, $c^T \Sigma c = (k(x_p, x_p) + k(x_o, x_o) - 2k(x_p, x_o)) - (k_n(\mathbf{x})c)^T (K_n + \sigma^2 I)^{-1} (k_n(\mathbf{x})c)$. Since $c^T \mu, c^T \Sigma c \in \mathbb{R}$, we have that $P(y_p - y_o > 0 | \mathcal{H}) = \Phi\left(\frac{c^T \mu}{\sqrt{c^T \Sigma c}}\right)$ (here $\Phi(\cdot)$ denotes the CDF of the standard normal). Therefore, the acceptance probability reduces to

$$\alpha = \min \left\{ 1, \frac{\Phi\left(\frac{c^T \mu}{\sqrt{c^T \Sigma c}}\right) q(x_p|x_o)}{(1 - \Phi\left(\frac{c^T \mu}{\sqrt{c^T \Sigma c}}\right)) q(x_o|x_p)} \right\}. \quad (2)$$

A demonstration of MCMC-BO with MH is shown in Algorithm 2. MCMC-BO prepares a batch of m candidate points x_{tk}^i each round either from direct discretization or points to be executed from other algorithms, where t stands for the round number, $k = 1, 2, 3, \dots, m$, and i stands for transition times. Then with a proposed transition kernel, often defined using Brownian motion $x^{i+1} - x^i \sim \mathcal{N}(0, \sigma)$, we generate m pairs of points x_{tk}^i, x_{tk}^{i+1} . We accept or decline the transition with the ratio in Equation 2. The random walk of Markov chain enables dense discretization of continuous space on more optimal regions, as illustrated in Fig 1.

The generality of our framework also extends to other sampling methods such as LD, where the samples transit following Equation ?? . Previously, MH aims to sample from the intractable distribution corresponding to our acquisition approach by simplifying the ratio between two discretization points. Denoting the time-varying density $p_t(x) \propto \mathbb{P}(f_t(x) \geq f_t(x'), \forall x' \in D_t)$, LD updates requires computing $\nabla \log p_t(x)$.

Similar to MH's acceptance ratio, we propose a simple estimate using the derivative of the GP posterior's mean and covariance.

In our proposed simplification for MCMC-BO, we utilize MH to calculate the ratio between the "winning probabilities" of two discretization points (the probability of being arg max). This calculation is done specifically with respect to the posterior distribution on those two points, rather

Algorithm 2: [MCMC routine] with Metropolis-Hastings

Input: GP posterior $\text{GP}(\mu_t(x), k_t(x, x'))$,
point x_o , proposal distribution
 $q(\cdot | x) := \mathcal{N}(0, \sigma)$ with parameter
 $\Theta = \{\sigma\}$.

- 1 Sample $u \sim \text{Unif}[0, 1]$
 - 2 Sample $x_p \sim x_o + \mathcal{N}(0, \Theta)$
 - 3 **if** $u \geq \min \left\{ 1, \frac{\mathbb{P}(f_t(x_p) \geq f_t(x_o))q(x_p|x_o)}{\mathbb{P}(f_t(x_o) \geq f_t(x_p))q(x_o|x_p)} \right\}$
 then
 - 4 $x_p \leftarrow x_o$ // Reject the transition
 - 5 **end**
-

Algorithm 3: [MCMC routine] with Langevin dynamics

Input: GP posterior $\text{GP}(\mu_t(x), k_t(x, x'))$,
point x , $\Theta = \{\epsilon\}$ as Langevin
transition step.

- 1 Sample $z \sim \mathcal{N}(0, 1)$
 - 2 Estimate $\nabla \log p_t(x)$ using (3)
 - 3 $x \leftarrow x + \epsilon \cdot \nabla \log p_t(x) + \sqrt{2\epsilon} \cdot z$
-

than the entire domain. The term $\nabla \log p_t(x)$ easily incorporates this simplified ratio as: $\frac{\partial \log p_t(x)}{\partial x_i} = \frac{\frac{\partial p_t(x)}{\partial x_i}}{p_t(x)} = \lim_{h \rightarrow 0} \frac{p_t(x+e_i h) - p_t(x)}{h p_t(x)} = \lim_{h \rightarrow 0} \frac{1}{h} \left(\frac{p_t(x+e_i h)}{p_t(x)} - 1 \right)$. Thus, the approximation of the log-likelihood in LD has the following equation:

$$\frac{\partial \log p_t(x)}{\partial x_i} \approx \lim_{h \rightarrow 0} \frac{1}{h} \left(\frac{\mathbb{P}(f_t(x+e_i h) > f_t(x))}{\mathbb{P}(f_t(x) > f_t(x+e_i h))} - 1 \right) \approx \frac{1}{h} \left(\frac{\mathbf{p}_i(x, h)}{1 - \mathbf{p}_i(x, h)} - 1 \right), \quad (3)$$

where $\mathbf{p}_i(x, h) = \mathbb{P}(f_t(x+e_i h) > f_t(x)) = \Phi\left(\frac{c^T \mu}{\sqrt{c^T \Sigma c}}\right)$. As in MH, this involves calculating the cumulative distribution function (CDF) of a bivariate Gaussian, with the difference being that $x_p \leftarrow x + e_i h$ and $x_o \leftarrow x$. Using this numerical differentiation of the likelihood ratio, we obtain the gradient of $\log p_t(x)$ and the LD form of MCMC-BO.

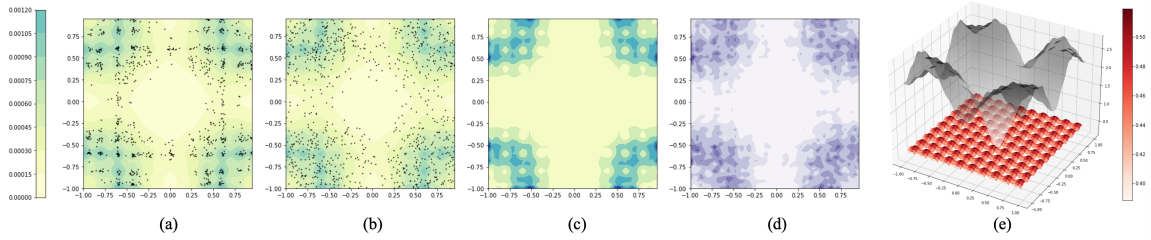


Figure 2: The figures are constructed from a 50×50 discretization of $D = [-1, 1]^2$. **(a)(b)** The stationary distribution achieved by the MH and Langevin version of MCMC-BO, respectively, and the congregated points obtained after convergence of the transition process from current GP information. **(c)** TS distribution simulated using Monte Carlo. **(d)** standard deviation of TS distribution over 10 trials of 10^6 samples. **(e)** GP posterior with surfaces being μ and Σ on which MCMC-BO transitions are performed.

5. Convergence Guarantee

We provide convergence guarantees to our proposed algorithmic framework and derive the subsequent regret bounds. Our proposed algorithm removes the dependence on the size of the discretization set; whereas the original GP-TS setting requires that to achieve the theoretical upper bound on regret. As we only track a batch of m points at a time, the algorithm complexity for MCMC-BO only depends on m and does not scale with the size of the discretization set $|D_t|$.

Notation and assumption. Let $\{x_t\}_{t=1}^\infty$ be an \mathbb{R}^d -valued discrete-time stochastic process adapted to filtration $\{\mathcal{F}_t\}_{t=0}^\infty$. Denote $x^* = \operatorname{argmax}_{x \in D} f(x)$, then the cumulative regret over the horizon T is $R_T = \sum_{t=1}^T r_t$, where $r_t = f(x^*) - f(x_t)$. We also follow the convention and denote the maximum information gain at time t as γ_t . Further notation is attached in Appendix B.

Leveraging the properties of Reproducing Kernel Hilbert Space (RKHS), we make a common assumption as in Chowdhury and Gopalan (2017), that f is Lipschitz constant with constant $B \cdot L$, where L is a constant associated with Kernel $k(\cdot, \cdot)$.¹ In our analysis, we work with a discretization D_t where $|D_t|$ is finite. We denote $[x]_t := \arg \min_{x' \in D_t} \|x - x'\|_2$. For convex and compact domain $D \subset [0, r]^d$, the size constraint $|D_t| = (BLrdt^2)^d$ guarantees that for every $x \in D$, $|f(x) - f([x]_t)| \leq \|f\|_k L \|x - [x]_t\|_1 \leq 1/t^2$. Note that the size of the discretization is only

1. Further detail of assumption can be seen in Appendix B

needed in the theoretical analysis and influences the regret bound, but the algorithm complexity is independent of it. Additionally, the transition kernel of MCMC is defined as $P(x, y) = Q(x, y)\alpha(x, y), \forall x, y \in D_t$, where Q is the proposed distribution and α is the acceptance rate.

5.1. Overview of Proof

First, we show that the proposed approximated ratio of TS has a stationary distribution. Both MCMC algorithm (MH and LD) can converge to the stationary distribution with certain steps. Then the stationary distribution, though deviated from that of Thompson sampling, still inherits its good properties. Thus, by defining a benign set, we can prove that the probability that the sampling points located in the benign set grows with $O(1 - \frac{1}{T^2})$. An overview is presented here, and the detailed proof is deferred to Appendix C.

Lemma 1 *The proposed approximated posterior $P(x, \cdot)$, according to Alg.2 and Equation.2 does not yield a reversible Markov chain, but it still has a stationary distribution $\pi(x)$.*

In Fig 2 we demonstrate a 2D version of the stationary distribution of MCMC-BO. Note that the acceptance probability our algorithm proposed is inspired by TS but does not lead to the same exact stationary distribution of TS. Furthermore, we show that the approximated probability $P(x, \cdot)$ converges to the stationary distribution. In Sec 6 we also show that within a few hundred steps, the effect is significant enough. To obtain a bound on the overall regret, we begin by decomposing the instantaneous regret r_t into two parts. From the choice of discretization sets D_t , r_t depends on $[x^*]_t$, which is the closest point to x^* in D_t . Therefore, we have that $r_t = (f(x^*) - f([x^*]_t)) + (f([x^*]_t) - f(x_t))$, where the difference $f(x^*) - f([x^*]_t)$ is bounded by $1/t^2$ by the regularity assumption. We proceed to bound $f([x^*]_t) - f(x_t)$, which depends on the selected action x_t .

Regret increases when the approximation deviates from the underlying f . Thus, we hope that the function values of selected points x_t and the approximation at each time step t are not too far away. We define an event $E^f(t)$ and a benign set G_t in Appendix C for the convenience of Lemma 2. The benign points set bound the difference between function value and at the $t - 1$'s round GP's mean value.

Lemma 2 *For any filtration such that $E^f(t)$ is true, the transition probability from $[x^*]_t$ to any malignant point is bounded by*

$$P([x^*]_t, x') \leq \exp(-(\frac{c_t}{2} - v_t)^2)(1 - t^2), \quad \forall x' \in D_t \setminus G_t \quad (4)$$

Lemma 2 shows that transiting on the stationary distribution $\pi(x)$, the probability of choosing a malignant points is small. With Lemma 2, we can further show that the probability of playing any action from $D_t \setminus G_t$ is small; the regret from undesirable action is then bounded with Theorem 5 in Appendix B. By giving the bound of the transition probability between the best point and the malignant points, the total regret can then be bounded in terms of standard deviation of the chosen actions $\sigma_{t-1}(x_t)$. With the information theoretic lemma (Lemma 3 in Appendix) we can then upper bound $\sum_{t=1}^T \sigma_{t-1}(x_t)$ with $O(\sqrt{T\gamma_T})$. As the difference $f([x^*]_t) - f(x_t)$ in the decomposition of r_t is dependent on $\sigma_{t-1}(x_t)$, we finally arrive at the regret bound.

Theorem 3 *Let $\delta \in (0, 1)$, $D \subset [0, r]^d$ be compact and convex, $\|f\|_k \leq B, \{\varepsilon_t\}_t$ a conditionally R -sub-Gaussian sequence. For any T , we let $N(T)$ to be the transition number for each round.*

Running MCMC-BO on a function f lying in the RKHS $H_k(D)$ for $N(T)$ transitions per round and with decision sets D_t chosen as above, with probability at least $1 - \delta$, the regret of MCMC-BO satisfies $R_T = O\left(\sqrt{(\gamma_T + \ln(2/\delta)) d \ln(BdT)} \left(\sqrt{T\gamma_T} + B\sqrt{T \ln(2/\delta)}\right)\right)$.

6. Experiments

In this section, we show the superiority of MCMC-BO on both high-dimensional synthetic functions and Mujoco tasks. We also provide the ablation study for the algorithm in Table 1 in Appendix. We further validate that our proposed method itself is an effective GP-based BO algorithm as other MAB algorithms. We compare MCMC-BO to state-of-the-art baselines of high-dimensional BO algorithms and EA algorithms.

For all BO algorithms, we utilize Thompson sampling to sample batches in each iteration and employ a scrambled Sobolev sequence to discretize the continuous search domain. The performance of MCMC-BO is evaluated against TuRBO and LA-MCTS, which serve as its BO components. The performance figures illustrate the mean performance of the algorithms with one standard error.

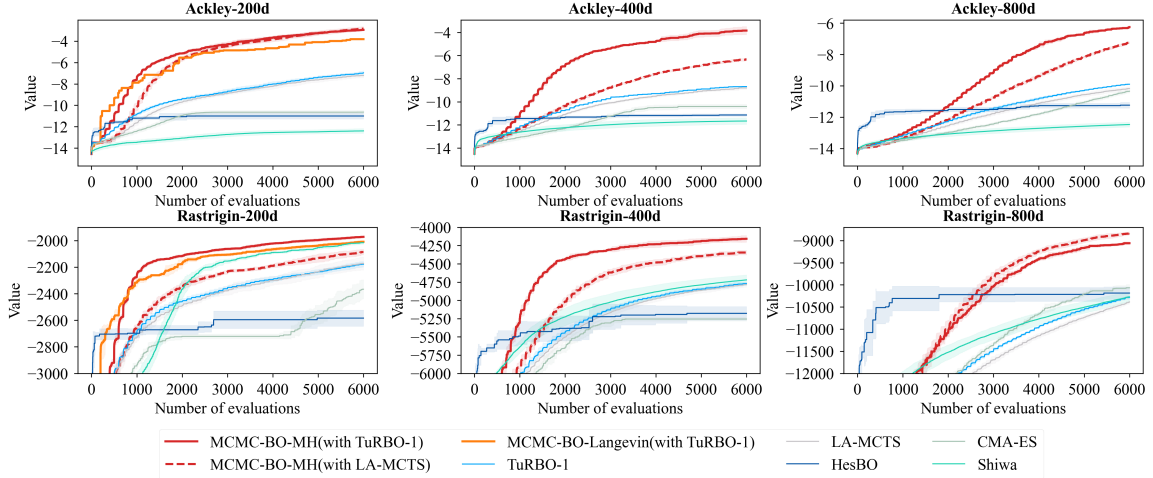


Figure 3: Optimization over high-dimensional synthetic functions.

6.1. High-Dimensional Synthetic Functions

We selected two popular synthetic problems, Ackley and Rastrigin, to optimize. For the function dimension, we chose $d = 200, 400$, and 800 to evaluate the performance in high-dimensional problems. We set the transition number to be the same as the function dimension. However, for the Langevin version of MCMC-BO, we ran only the $200d$ functions due to the lack of backward information for the posterior in the Gpytorch framework. All problems started with 200 initial points, and a batch of 100 points was sampled in each iteration.

Fig 3 suggests that MCMC-BO consistently outperforms other baselines on all functions. In higher dimensions $d = 800$, uniform discretization can not support good exploitation in such dimension, and TuRBO and LA-MCTS degenerate to the same level of performance as EA algorithms.

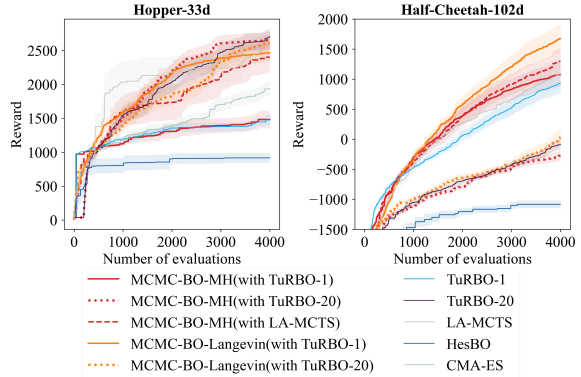


Figure 4: Mujoco locomotion tasks.

MCMC-BO achieves the best performance in all selected dimensions by allocating limited action points to more promising regions.

6.2. Mujoco Locomotion Task

Mujoco locomotion tasks are widely used benchmarks for reinforcement learning algorithms (Todorov et al., 2012). In our evaluation, we focus on Hopper and Half-Cheetah task, which have state spaces of dimensions 33 and 102, respectively. To assess the performance of the sampling-based algorithms, we optimize a linear policy: $\mathbf{a} = \mathbf{W}\mathbf{s}$ (Mania et al., 2018), where the elements of the parameter matrix \mathbf{W} are continuous and range from $[-1, 1]$. The reward is computed over 10 episodes for each policy proposal. Both tasks start with 200 initial points and sample a batch of 50 points in each iteration. We set the transition number to 200 on both tasks. Fig 6.1 shows the optimization performance of all algorithms. In the Hopper task, MCMC-BO with TuRBO-20 algorithm converges faster than the original TuRBO-20. In higher dimensional Half-Cheetah task, MCMC-BO with TuRBO-1 still outperforms other baselines.

6.3. Performance on low-dimensional problems

We compare MCMC-BO with two popular BO algorithms, GP-TS and GP-UCB, on 1-2d synthetic functions. Figure 5 numerically demonstrates the convergence of MCMC-BO. We depict the distribution over the optimal value encoded as a violin plot, with horizontal bars 20% quantiles. GP-UCB only utilizes the diagonal information of the covariance matrix of $f(x_i)$, whereas TS uses the complete matrix information at once. As a compromise, MCMC-BO employs a divide-and-conquer strategy by using a 2×2 sub-matrix at a time and still achieving an approximate stationary distribution of TS.

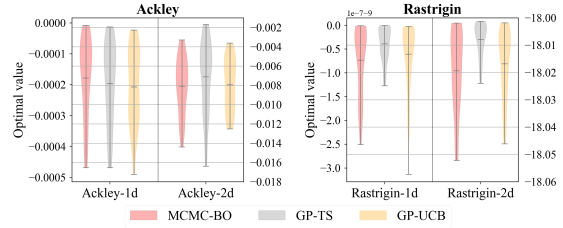


Figure 5: We demonstrate the convergence of MCMC-BO compared with others on 1d and 2d functions.

7. Conclusion and Future Work

Sequential optimization in high-dimensional spaces has a profound impact on machine learning. In this paper, we propose MCMC-BO as a solution to improve the sample efficiency of high-dimensional BO. Our algorithm MCMC-BO offers versatile transitions to promising regions instead of maintaining a huge candidate set. MCMC-BO can be viewed as a GP-based bandit algorithm that yields an effective approximation for the induced TS distribution without the need to invert matrices of thousands of dimensions. We derive the regret bound of MCMC-BO under high-dimensional cases without memory overuse. We also conduct comprehensive evaluations to show that MCMC-BO can improve on existing popular high-dimensional BO baselines.

In future research, we aim to develop parallel computation mechanisms to further enhance computational speed. Additionally, implementing analytical backward computations holds the potential for significant acceleration. It is worth noting that MCMC-BO can transit in different irregular spaces. We also look forward to combining MCMC-BO with more complex space partition algorithms.

APPENDIX

Appendix A. Algorithm

We present a baseline of MCMC-BO compatible with any other existing probabilistic BO algorithms over continuous domain. Various BO algorithms propose different acquisition functions, either implicitly or explicitly. MCMC-BO is able to take points proposed by BO methods into the MCMC transition step and subsequently find a more promising region.

Algorithm 4: MCMC-BO with existing BO algorithms

Input: Initial dataset \mathcal{D}_0 , Acquisition function A , batch size m , MCMC transition number N , transition noise σ

```

1 for  $t = 0, 1, \dots$  do
2   Update posterior distribution on  $f$  using  $\mathcal{D}_t$ 
3   Create discretized candidate points  $\mathbf{x}_{\text{cand}}$  from continuous search domain
4   Propose a batch  $\mathbf{x}_t^0 = \operatorname{argmax}_{\mathbf{x}_{\text{cand}}} A(\mathbf{x}_{\text{cand}} | \mathcal{D}_t)$ 
   // MCMC transition start
5   for  $i = 0$  to  $N - 1$  do
6     for  $k = 0$  to  $m - 1$  do
7       Sample  $u \sim \text{Unif}[0, 1]$ 
8       Sample  $x_{tk}^{i+1} \sim x_{tk}^i + \mathcal{N}(0, \sigma)$ 
9       if  $u \geq \min \left\{ 1, \frac{P(f_t(x_{tk}^{i+1}) \geq f_t(x_{tk}^i) | q(x_{tk}^{i+1} | x_{tk}^i))}{P(f_t(x_{tk}^i) \geq f_t(x_{tk}^{i+1}) | q(x_{tk}^i | x_{tk}^{i+1}))} \right\}$  then
10        |  $x_{tk}^{i+1} \leftarrow x_{tk}^i$  // Reject the transition
11        | else
12        | // Accept the transition (no action needed if accepted)
13        | end
14      end
15    end
   // MCMC transition end
16   Observe  $\mathbf{y}_t = f(\mathbf{x}_t^N)$ 
17    $\mathcal{D}_{t+1} \leftarrow \mathcal{D}_t \cup (\mathbf{x}_t^N, \mathbf{y}_t)$ 
18 end

```

Based on this algorithmic framework, we can also derive the Gibbs sampling counterpart of MCMC-BO where we use Gibbs sampling procedures in place of MH. In fact, Gibbs sampling can be viewed as a special case of MH where the acceptance probability equals 1. Gibbs sampling transits the point x with respect to the conditional distribution of each input dimension one at a time, with $Q(x'_d, x_{-d} | x) = x'_d \sim U(x_d)$ where x_d, x^{-d} stands for the d^{th} entry of the vector x and x^{-d} stands for the remaining $d - 1$ entries. The acceptance probability becomes $\alpha = \frac{p(x'_d, x_{-d})Q(x_d, x_{-d} | x'_d, x_{-d})}{p(x_d, x_{-d})Q(x'_d, x_{-d} | x_d, x_{-d})} = 1$. Gibbs sampling algorithm combined with MCMC-BO is a more suitable fit for vanilla Bayesian optimization with randomized initialization techniques such as Sobolev sequences, as it transitions candidate points on one dimension at each round while utilizing Thompson sampling.

Appendix B. Detailed Theoretical Analysis

Notation and assumption. Let $\{x_t\}_{t=1}^\infty$ be an \mathbb{R}^d -valued discrete-time stochastic process adapted to filtration $\{\mathcal{F}_t\}_{t=0}^\infty$, i.e., x_t is \mathcal{F}_{t-1} measurable, $\forall t \geq 1$. Let $\{\varepsilon_t\}_{t=1}^\infty$ be a real-valued stochastic process such that for some $R \geq 0$ and for all $t \geq 1$, ε_t is \mathcal{F}_t -measurable and R -sub-Gaussian conditionally on \mathcal{F}_{t-1} . We use the notion of regret as the evaluation metric for the performance of our algorithm. Denote $x^* = \arg\max_{x \in D} f(x)$, then the cumulative regret over the horizon T is $R_T = \sum_{t=1}^T r_t$, where $r_t = f(x^*) - f(x_t)$. The sample from the GP posterior at time t is denoted by f_t , and the corresponding mean and covariance (on the discretization set) μ_t, Σ_t . We also follow the convention and denote the maximum information gain at time t as γ_t .

We first give a formal definition of Markov chain and the transition kernel.

Definition 4 (Transition kernel) Consider two measure spaces (Ω, \mathcal{A}) and (Ω', \mathcal{A}') . The function $K : \Omega \times \mathcal{A}' \rightarrow [0, \infty)$ is a kernel from (Ω, \mathcal{A}) to (Ω', \mathcal{A}') if

- $K(\omega, \cdot) : \mathcal{A}' \rightarrow [0, \infty)$ is a measure on (Ω', \mathcal{A}') .
- $K(\cdot, \mathcal{A}') : \Omega \rightarrow [0, \infty)$ is \mathcal{A} -measurable.

If $K(\omega, \Omega') = 1$, the kernel is a Markov Kernel.

We impose standard smoothness assumptions on f in the Reproducing Kernel Hilbert Space (RKHS). Let $k : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$ be a symmetric, positive-semidefinite kernel. For $f \in H_k(D)$, $\|f\|_k \leq B$, where $H_k(D)$ is the induced RKHS. This assumption is equivalent to saying f is Lipschitz with constant $B \cdot L$, where $L = \sup_{x \in D} \sup_{j \in [d]} \left(\frac{\partial^2 k(p, q)}{\partial p_j \partial q_j} \Big|_{p=q=x} \right)^{1/2}$ (Chowdhury and Gopalan, 2017). In our analysis, we work with a discretization D_t , and the transition kernel is supported on the probability space $(D_t, \mathcal{B}(D_t))$, where $|D_t|$ is finite. We denote $[x]_t := \arg \min_{x' \in D_t} \|x - x'\|_2$. For convex and compact domain $D \subset [0, r]^d$, the size constraint $|D_t| = (BLrdt^2)^d$ guarantees that for every $x \in D$, $|f(x) - f([x]_t)| \leq \|f\|_k L \|x - [x]_t\|_1 \leq 1/t^2$. Therefore, we can bound the deviations of available points from this time-varying discretization to the true optimum which may not be obtainable due to such discretization. Note that the size of the discretization is only needed in the theoretical analysis and influences the regret bound, but the algorithm complexity is independent of it.

We also impose smoothness assumptions on f in the Reproducing Kernel Hilbert Space (RKHS) such that for $f \in H_k(D)$, $\|f\|_k \leq B$. Here $H_k(D)$ is the corresponding RKHS, note that the RKHS norm satisfies $f \in H_k(D)$ if and only if $\|f\|_k < \infty$. Moreover, f should be Lipschitz continuous with constant $B \cdot L$. In our analysis, we work with a discretization D_t of time-varying size as a necessary tool for bounding the regret. However, such discretization is only an analytical tool rather than a practical requirement. we choose a compact and convex domain $D \subset [0, r]^d$ and discretization D_t with size $|D_t| = (BLrdt^2)^d$ such that $\|x - [x]_t\|_1 \leq rd/BLrdt^2 = 1/BLt^2$ for all $x \in D$, where $L = \sup_{x \in D} \sup_{j \in [d]} \left(\frac{\partial^2 k(p, q)}{\partial p_j \partial q_j} \Big|_{p=q=x} \right)^{1/2}$. This implies, for every $x \in D$, $|f(x) - f([x]_t)| \leq \|f\|_k L \|x - [x]_t\|_1 \leq 1/t^2$,

Definition 5

1. A Markov chain with n -step transition probability $P^n(x, A) = P(\Phi_n \in A \mid \Phi_0 = x)$ is φ -**irreducible** if there exists a measure φ such that

$$\varphi(A) > 0 \Rightarrow \sum_n P^n(x, A) > 0, \quad \text{for } x \in X.$$

2. For a φ -irreducible chain, any set A with $\varphi(A) > 0$ contains a **small set** C such that for $\delta > 0$, $n > 0$ and probability measure ν concentrated on C ,

$$P^n(x, \cdot) \geq \delta \nu(\cdot), \quad \text{for } x \in C. \quad (5)$$

3. The chain is **aperiodic** if for some small set C with $\varphi(C) > 0$, the greatest common divisor of all the n such that (5) is true is 1.
4. The chain Φ with transition probability $P^n(x, \cdot)$ and an invariant distribution ϕ for Φ is **uniformly ergodic** if there exists a sequence $\tau(n)$ converging to 0 such that for all x , $\|P^n(x, \cdot) - \pi\| \leq \tau(n)$.

Theorem 6 (Mengersen and Tweedie, 1996) For any Markov chain, the following are equivalent:

1. The chain is aperiodic and the following condition holds: there is a probability measure ϕ on $\mathcal{B}(X)$ and $\epsilon > 1$, $\delta > 0$, $m \in \mathbb{Z}_{>0}$ such that when $\phi(A) > \epsilon$, $\inf_{x \in X} P^m(x, A) > \delta$.
2. For some small set C we have $\sup_{x \in X} E_x[\tau_C] < \infty$, where τ_C is the first return time to C .
3. Φ is uniformly ergodic.

When any of the above criteria are satisfied, for any x we can bound the rate of convergence by

$$\|P^n(x, \cdot) - \pi\| \leq (1 - \delta)^{\lfloor n/m \rfloor}.$$

Lemma 7 A chain Φ is π -irreducible if $\pi(y) > 0 \Rightarrow q(x, y) > 0$, for $x \in X$. The MH chain Φ is aperiodic and every compact set C with $\mu(C) > 0$ is small if $\pi(x)$ and $q(y, x)$ are positive and continuous for all x, y .

Proof See (Mengersen and Tweedie, 1996). ■

Lemma 8 The MCMC-BO's Markov chain is **aperiodic**, **acyclic** and **p-irreducible**, it has stationary distribution. And any state x has a finite expectation of first return time, i.e. $E_x[\tau_x] < \infty$ and $E_x[\tau_C] < \infty$.

Lemma 1 The proposed approximated posterior $P(x, \cdot)$, according to Alg 2 and Equation 2 does not yield a reversible Markov chain, but it still has a stationary distribution $\pi(x)$.

Proof Intuitively, MH algorithm leverages the acceptance ratio of stationary distribution. For a symmetrical proposal distribution, we have that $q(x_o, x_p) = q(x_p, x_o) = \mathcal{Q}(x_o - x_p)$ where $x_o, x_p \in D_t$. Therefore, it follows that whether the ratio $\frac{P(f_t(x_p) \geq f_t(x_o))q(x_p|x_o)}{P(f_t(x_o) \geq f_t(x_p))q(x_o|x_p)}$ is greater than one solely depends on the value $\mu(x_o) - \mu(x_p)$. Therefore, we can rank the distribution according to $\mu(x)$

given by the posterior. For any three point x_i, x_j, x_k satisfying $\mu(x_i) > \mu(x_j) > \mu(x_k)$ if the MH chain is reversible, then $\pi(x_i)P_{ij} = \pi(x_j)P_{ji}$ and should satisfy that

$$\frac{\pi(x_i)}{\pi(x_j)} \frac{\pi(x_j)}{\pi(x_k)} \frac{\pi(x_k)}{\pi(x_i)} = \frac{P_{ij}}{P_{ji}} \frac{P_{kj}}{P_{jk}} \frac{P_{ki}}{P_{ik}} = 1$$

From the given rank of index i, j , and k , the probabilities $P(f_t(x_i) \geq f_t(x_j)), P(f_t(x_j) \geq f_t(x_k)), P(f_t(x_i) \geq f_t(x_k))$ are greater than $\frac{1}{2}$. Therefore, $\frac{P_{ij}}{P_{ji}} \frac{P_{kj}}{P_{jk}} \frac{P_{ki}}{P_{ik}} = \frac{P(f_t(x_j > x_i))}{P(f_t(x_i > x_j))} \frac{P(f_t(x_k > x_j))}{P(f_t(x_j > x_k))}$. $1 < 1$. Contradicts with the assumption. But as discretization set D_t , it has finite state number N_t . From Lemma 8 and Lemma 7, it has a stationary distribution. \blacksquare

Lemma 2 For any stationary distribution from the MCMC-BO's chain $\pi(x)$, $\exists n_\pi$ such that $\forall n \geq n_\pi$ we can reach a close distribution $\|P^n(x, \cdot) - \pi\| \leq \epsilon$.

Proof This follows Lemma 1 and Theorem 6. \blacksquare

Theorem 9

(Chowdhury and Gopalan, 2017) Recall that the covariance function $k : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$ is assumed to be a symmetric, positive-semidefinite kernel with associated feature map $\varphi : \mathbb{R}^d \rightarrow H_k$ and RKHS H_k .

Let $0 < \delta \leq 1$. For a given $\eta > 0$, with probability at least $1 - \delta$, the following holds: $\forall t > 1$,

$$\|\varepsilon_{1:t}\|_{((K_t + \eta I)^{-1} + I)^{-1}}^2 \leq 2R^2 \ln \frac{\sqrt{\det((1 + \eta)I + K_t)}}{\delta}.$$

Letting $S_t = \sum_{s=1}^t \varepsilon_s \varphi(x_s)$ and define the (possibly infinite dimensional) matrix $V_t = I + \sum_{s=1}^t \varphi(x_s) \varphi(x_s)^T$, we have, whenever K_t is positive definite, that

$$\|\varepsilon_{1:t}\|_{(K_t^{-1} + I)^{-1}} = \|S_t\|_{V_t^{-1}},$$

where $\|S_t\|_{V_t^{-1}} := \left\| V_t^{-1/2} S_t \right\|_{H_k}$ denotes the norm of $V_t^{-1/2} S_t$ in the RKHS H_k .

Here, K_t denotes the $t \times t$ matrix $K_t(i, j) = k(x_i, x_j)$, $1 \leq i, j \leq t$ and for any $x \in \mathbb{R}^t$ and $A \in \mathbb{R}^{t \times t}$, $\|x\|_A := \sqrt{x^T A x}$. Moreover, if K_t is positive definite $\forall t \geq 1$ with probability 1, then the conclusion above holds with $\eta = 0$.

Lemma 10 For every $t \geq 0$, the maximum information gain γ_t , for the points chosen by MCMC-BO satisfy, almost surely, the following :

$$\begin{aligned} \gamma_t &\geq \frac{1}{2} \ln (\det (I + \lambda^{-1} K_t)) \\ \gamma_t &\geq \frac{1}{2} \sum_{s=1}^t \ln (1 + \lambda^{-1} \sigma_{s-1}^2(x_s)). \end{aligned}$$

Proof The proof follows from (Chowdhury and Gopalan, 2017), Lemma 3. \blacksquare

Lemma 11 (*Chowdhury and Gopalan, 2017*) *Let x_1, \dots, x_t be the points selected by the algorithms. The sum of predictive standard deviation at those points can be expressed in terms of the maximum information gain. More precisely,*

$$\sum_{t=1}^T \sigma_{t-1}(x_t) \leq \sqrt{4(T+2)\gamma_T}.$$

Lemma 12 (*Gaussian Anti-concentration*) *For a Gaussian random variable X with mean μ and standard deviation σ , for any $\beta > 0$,*

$$\mathbb{P} \left[\frac{X - \mu}{\sigma} > \beta \right] \geq \frac{e^{-\beta^2}}{4\sqrt{\pi}\beta}.$$

Definition 13 *Define $E^f(t)$ as the event that such that for all $x \in D$, $|\mu_{t-1}(x) - f(x)| \leq v_t \sigma_{t-1}(x)$ where $v_t = B + R\sqrt{2(\gamma_{t-1} + 1 + \ln(2/\delta))}$.*

Definition 14 (*Benign points*) *Define the set of points with better estimated mean in discretization D_t at round t as $G_t := \{x \in D_t : \mu_{t-1}(x) \geq f([x^*]_t) - c_t \sigma_{t-1}(x)\}$, where $\tilde{c}_t = \sqrt{4 \ln t + 4d \ln(BLrdt^2)}$ and $c_t = v_t(1 + \tilde{c}_t)$.*

Definition 15 *Define the set of saturated points S_t in discretization D_t at round t as*

$$S_t := \{x \in D_t : \Delta_t(x) > c_t \sigma_{t-1}(x)\}$$

where $\Delta_t(x) := f([x^*]_t) - f(x)$, the difference between function values at the closest point to x^* in D_t and at x . Clearly $\Delta_t([x^*]_t) = 0$ for all t , and hence $[x^*]_t \in D_t$ is unsaturated at every t .

Theorem 16 (*Chowdhury and Gopalan, 2017*) *Under the same hypotheses as those of Theorem 9, let $D \subset \mathbb{R}^d$, and $f : D \rightarrow \mathbb{R}$ be a member of the RKHS of real-valued functions on D with kernel k , with RKHS norm bounded by B . Then, with probability at least $1 - \delta$, the following holds for all $x \in D$ and $t \geq 1$: $|\mu_{t-1}(x) - f(x)| \leq \left(B + R\sqrt{2(\gamma_{t-1} + 1 + \ln(1/\delta))} \right) \sigma_{t-1}(x)$, where γ_{t-1} is the maximum information gain after $t - 1$ rounds and $\mu_{t-1}(x), \sigma_{t-1}^2(x)$ are mean and variance of posterior distribution defined as in Equation 2, 3, 4, with λ set to $1 + \eta$ and $\eta = 2/T$.*

Lemma 17 *For any \mathcal{F}'_{t-1} such that $E^f(t)$ is true. We consider one of the benign points in D_t with smallest $\sigma_{t-1}(x)$, i.e. $\bar{x}_t = \arg\min_{x \in G_t} \sigma_{t-1}(x)$. We have that the expectation of the variance of \bar{x}_t satisfies that*

$$\mathbb{E} [\sigma_{t-1}(x_t) \mid \mathcal{F}'_{t-1}] \geq \sigma_{t-1}(\bar{x}_t) \left(\frac{1}{4e\sqrt{\pi}} - 1/t^2 \right) \quad (6)$$

Lemma 18 *Given any $\delta \in (0, 1)$, $\mathbb{P} [E^f(t), \forall t \geq 1] \geq 1 - \delta/2$ and for all possible history \mathcal{F}'_{t-1} , $\mathbb{P} [E^{f_t}(t) \mid \mathcal{F}'_{t-1}] \geq 1 - 1/t^2$.*

Proof The proof follows from (*Chowdhury and Gopalan, 2017*), Lemma 5. ■

Theorem 19 Consider the stationary distribution as π and the proposal distribution as Q and divide the discretization set D_t into three subgroups D_t^1, D_t^2, D_t^3 satisfying that for any $x_{D_1} \in D_t^1, x_{D_2} \in D_t^2, x_{D_3} \in D_t^3, \mu_{t-1}(x_{D_1}) \geq \mu_{t-1}(x_{D_2}) \geq \mu_{t-1}(x_{D_3})$. Further denote $\pi_i = \sum_{x \in D_i} P(x)$. For the proposal distribution Q , we have that

$$\pi_3 \leq o(1)$$

Proof

Suppose that $\pi = [\pi_1, \pi_2, \pi_3]$ is the stationary distribution of the three subgroups, then we must have that

$$(\pi_1 \quad \pi_2 \quad \pi_3) \left(\begin{pmatrix} Q_{11} & Q_{12} & Q_{13} \\ Q_{21} & Q_{22} & Q_{23} \\ Q_{31} & Q_{32} & Q_{33} \end{pmatrix} \cdot \begin{pmatrix} 1 & \alpha_{12} & \alpha_{13} \\ 1 & 1 & \alpha_{23} \\ 1 & 1 & 1 \end{pmatrix} \right) = (\pi_1 \quad \pi_2 \quad \pi_3), \quad (7)$$

where α_{ij} denotes the acceptance probability of transferring group i to group j . By assumption of $\mu_{t-1}(x_{D_i})$, we known that for $i \geq j, \alpha_{ij} = 1$. Thus, Equation (7) gives

$$\pi_3 = \pi_1 Q_{13} \alpha_{13} + \pi_2 Q_{23} \alpha_{23} + \pi_3 Q_{33}.$$

Substituting $\pi_1 = 1 - \pi_2 - \pi_3$, we have

$$\begin{aligned} \pi_3 &= (1 - \pi_2 - \pi_3) Q_{13} \alpha_{13} + \pi_2 Q_{23} \alpha_{23} + \pi_3 Q_{33} \\ \pi_3 (1 + Q_{13} \alpha_{13} - Q_{33}) &= Q_{13} \alpha_{13} + \pi_2 (Q_{23} \alpha_{23} - Q_{13} \alpha_{13}) \end{aligned}$$

Denote

$$\begin{aligned} (\alpha_{ij})_{\min} &:= \min_{x, x'} \alpha(x, x') \quad \text{s.t. } x \in D_t^i, x' \in D_t^j \\ (\alpha_{ij})_{\max} &:= \max_{x, x'} \alpha(x, x') \quad \text{s.t. } x \in D_t^i, x' \in D_t^j, \end{aligned}$$

and notice that $\forall i, j, \pi_2 = 1 - \pi_1 - \pi_3 \leq 1 - \pi_3$ and $\alpha_{ij} \leq 1$. We thus have

$$\begin{aligned} \pi_3 (1 + Q_{13} (\alpha_{13})_{\min} - Q_{33}) &\leq Q_{13} (\alpha_{13})_{\max} - \pi_2 Q_{13} (\alpha_{13})_{\min} + \pi_2 Q_{23} \alpha_{23} \\ &\leq Q_{13} (\alpha_{13})_{\max} - \pi_2 Q_{13} (\alpha_{13})_{\min} + (1 - \pi_3) Q_{23} \\ \pi_3 (1 + Q_{13} (\alpha_{13})_{\min} - Q_{33} + Q_{23}) &\leq Q_{13} (\alpha_{13})_{\max} = \pi_2 Q_{13} (\alpha_{13})_{\min} + Q_{23} \\ \pi_3 &\leq \frac{Q_{13} (\alpha_{13})_{\max} = \pi_2 Q_{13} (\alpha_{13})_{\min} + Q_{23}}{1 + Q_{13} (\alpha_{13})_{\min} - Q_{33} + Q_{23}} \end{aligned}$$

W.L.O.G., suppose that Q , the proposal distribution, is a uniform, then following the assumptions of GP-TS, we can assume that the size of D_t is approximately $c \cdot T^2$, for c a constant. Denote m as the number of elements in the support of π_1 , then we have

$$\pi_3 \leq \frac{Q_{13} \cdot \alpha_{13}}{1 - Q_{33} + Q_{13} \alpha_{13}} \leq \frac{Q_{13} \alpha_{13}}{1 - Q_{33}} = \frac{1 - \frac{m}{Ct^2}}{\frac{m}{Ct^2}} \cdot \alpha_{13}.$$

From TS and the posterior probability given by GP, we denote μ_i, σ_i corresponding to $\mu_{t-1}(x_i)$ and $\sigma_{t-1}(x_i)$ for ease of notation. The acceptance probability is given by

$$\alpha_{13} = \frac{\Phi\left(\frac{\mu_3 - \mu_1}{\sqrt{\sigma_1^2 + \sigma_3^2 - 2\rho\sigma_1\sigma_3}}\right)}{1 - \Phi\left(\frac{\mu_3 - \mu_1}{\sqrt{\sigma_1^2 + \sigma_3^2 - 2\rho\sigma_1\sigma_3}}\right)}$$

(here ρ denotes the correlation coefficient).

Suppose we could obtain a bound on $O(\gamma_T)$, i.e. we proceed by assuming $O(\gamma_T) \leq k_t$. Notice that since $\Phi\left(\frac{\mu_3 - \mu_1}{\sqrt{\sigma_1^2 + \sigma_3^2 - 2\rho\sigma_1\sigma_3}}\right) \leq \Phi\left(\frac{\mu_3 - \mu_1}{\sigma_1 + \sigma_3}\right)$. Using the classical Hoeffding inequality, $\Phi(x) \leq \exp(-x^2/2)$. we thus have

$$\alpha_{13} \leq \frac{e^{\frac{\mu_3 - \mu_1}{\sigma_1 + \sigma_3}}}{1 - e^{\frac{\mu_3 - \mu_1}{\sigma_1 + \sigma_3}}}.$$

Suppose there exists k_t such that $\frac{\mu_3 - \mu_1}{\sigma_1 + \sigma_3} \leq -k_t$, then we can bound π_3 by

$$\begin{aligned} \pi_3 &\leq \frac{ct^2 - m}{m} \cdot \frac{e^{-k_t^2/2}}{1 - e^{-k_t^2/2}} \\ &= \frac{ct^2 - m}{m} \frac{1}{e^{k_t^2/2} - 1}, \end{aligned}$$

and

$$\lim_{t \rightarrow \infty} \frac{ct^2 - m}{m} \frac{1}{e^{k_t^2/2} - 1} \in o(1)$$

We then verify the existence of k_t such that $\frac{\mu_3 - \mu_1}{\sigma_1 + \sigma_2} \leq -k_t$ is correct, which is equivalent to verifying that

$$\frac{\mu_3 - \mu_1}{\sigma_1 + \sigma_2} \leq -k_t \iff \mu_3 - k_t\sigma_3 \leq \mu_1 - k_t\sigma_1.$$

Recall the definition of $E^f(t)$ and S_t so that for a point x , we have $|f(x) - \mu_{t-1}(x)| \leq v_t\sigma_{t-1}(x)$ and $f(x) \leq f(x^*) - c_t\sigma_{t-1}(x)$. We define a set W_t such that for $x \in W_t$, $\mu_{t-1}(x) \leq f(x^*) - c_t\sigma_{t-1}(x) + v_t\sigma_{t-1}(x) = f(x^*) - v_t\tilde{c}_t\sigma_{t-1}(x)$. Then, for $x_3 \in W_t$,

$$\mu_3 + k_t\sigma_3 \leq f(x^*) - v_t\tilde{c}_t\sigma_3 + k_t\sigma_3.$$

On the other hand, for a point x_1 satisfying

$$f(x) \geq f(x^*) - v_t\sigma_{t-1}(x),$$

we must have $x_1 \in D_t \setminus S_t$ (this probably should be move to definition of D_1). Then, by definition of S_t , we have that

$$\begin{aligned} \mu_1 + v_t\sigma_1 &\geq f(x^*) - v_t\sigma_1 \\ \mu_1 &\geq f(x^*) - 2v_t\sigma_1 \\ \Rightarrow \mu_1 - k_t\sigma_1 &\geq f(x^*) - 2v_t\sigma_1 - k_t\sigma_1. \end{aligned}$$

Then we can solve for k_t by aligning the two inequalities

$$\begin{aligned} \mu_3 + k_t\sigma_3 &\leq \mu_1 - k_t\sigma_1 \\ \Rightarrow f(x^*) - 2v_t\sigma_1 - k_t\sigma_1 &\geq f(x^*) - v_t\tilde{c}_t\sigma_3 + k_t\sigma_3 \\ \Rightarrow -2v_t\sigma_1 - k_t\sigma_1 &\geq v_t\tilde{c}_t\sigma_3 + k_t\sigma_3. \end{aligned}$$

Hence, the inequality is satisfied by any value of k_t satisfying

$$k_t \leq (\tilde{c}_t - 2)v_t.$$

We can of course take $k_t = (\tilde{c}_t - 2)v_t$. Since $\tilde{c}_t = \sqrt{4 \ln t + 2d \ln(BLr dt^2)}$ and $v_t = B + R\sqrt{2(\gamma_{t-1} + 1 + \ln(2/\delta))}$, and $\gamma_T \in O((\ln T)^{d+1})$, then we obtain the previous conclusion by substituting in $k_t \geq \sqrt{4d \ln(BLr dt^2)} \gamma_t \geq O(\sqrt{\ln(BLr dt^2)}((\ln T)^{d+1}))$.

$$\begin{aligned} \lim_{t \rightarrow \infty} \frac{Q_{13} \alpha_{13}}{1 - Q_{33}} &\leq \lim_{t \rightarrow \infty} (BLr)^{dt^{2d}} \frac{1}{e^{k_t^2/2} - 1} \\ &\leq \frac{(BLr)^{dt^{2d}}}{e^{\ln(BLr dt^2)(\ln T)^{d+1}}} \\ &= O\left(e^{2d \ln t - \ln(BLr dt^2)(\ln T)^{d+1}}\right) = o(1) \end{aligned} \tag{8}$$

And therefore we bound π_3 successfully. ■

Theorem 20 For any filtration \mathcal{F}'_{t-1} such that $E^f(t)$ is true,

$$\mathbb{P}[x_t \in D_t \setminus S_t \mid \mathcal{F}'_{t-1}] \geq p - 1/t^2 \tag{9}$$

Proof The proof follows from Lemma 17, Lemma 18 and Theorem 19. ■

Appendix C. Experiment Details

Results on three synthetic functions: Ackley (with domain $[-5, 10]^d$), Rastrigin and Levy (with domain $[-5, 5]^d$) were presented. The implementation and the hyperparameters used are detailed below.

C.1. Algorithm Implementation

MCMC-BO: We use Gpytorch to implement GP inference in MCMC-BO (Gardner et al., 2018). When combining MCMC-BO with TuRBO and LA-MCTS, we set transition noise as $0.008 * L$, where L is the size vector of trust region.

TuRBO: TuRBO is implemented based on tutorials from Botorch (Balandat et al., 2020). While the tutorial only has incomplete demonstration of TuRBO-1, we further modify the original code to TuRBO-M version. The hyperparameters are set as default setting in authors' reference implementation (Eriksson et al., 2019).

LA-MCTS: LA-MCTS, we refer to authors' reference implementations, and use TuRBO-1 as its local BO solver (Wang et al., 2020).

HesBO: For HesBO, we refer to authors' reference implementations (Nayebi et al., 2019). We transformed default GP component into Gpytorch version for faster inference speed on GPU. We set the embedding dimension to 20 for all tasks

CMA-ES: We use pycma² to implement CMA-ES, and use default setting except setting population size equals to batch size.

2. <https://github.com/CMA-ES/pycma>

Shiwa: We use Nevergrad³ to implement Shiwa, and use default setting to run experiments.

All algorithms begin with N_{init} initial oints from Latin hypercube design (McKay et al., 2000). For TuRBO-M, each trust region gets $\frac{N_{\text{init}}}{M}$ initial points, where M is the number of trust regions. In every iteration, we discretize the search domain to 5,000 points using scrambled Sobolev sequence. The results are averaged over

- Synthetic function: 5 trials.
- Mujoco functions: 30 trials.
- Ablation study: 5 trials.
- Low dimensional performance: 50 trials.

C.2. Mujoco Experiment

We use the mean and standard deviation from (Mania et al., 2018) to normalize observations of Mujoco task. While observations in Mujoco are noisier than synthetic functions, the results on all Mujoco functions are computed by 30 repeats.

C.3. Ablation Study

Table 1 shows the optimization performance under different transition numbers.

Function	MCMC transition number			
	20	100	300	800
Ackley-20d	-0.38 ± 0.21	-0.12 ± 0.06	-0.18 ± 0.02	-0.12 ± 0.14
Rastrigin-20d	-5.27 ± 2.50	-4.20 ± 2.23	-2.13 ± 2.01	-3.51 ± 3.20

Table 1: Ablation studies on hyper-parameters of MCMC-BO. The results are the means and standard deviations of the five replicate independent experiments on Ackley-20d and Rastrigin-20d when the transition noise is 0.001 (The result of the Rastrigin function is added by 180).

The result reveals that the performance of MCMC-BO is not sensitive to transition noise within a reasonable range and when the transition number is large enough. As the transition number increases, the performance steadily improves until convergence at a certain number.

Table 2 shows the optimization performance under different transition noise.

Function	MCMC transition noise				
	0.0001	0.0005	0.001	0.005	0.01
Ackley-20d	-0.51 ± 0.08	-0.25 ± 0.27	-0.18 ± 0.02	-0.17 ± 0.02	-0.22 ± 0.06
Rastrigin-20d	-5.41 ± 2.30	-4.59 ± 2.05	-2.13 ± 2.01	-2.49 ± 0.71	-4.17 ± 1.29

Table 2: Ablation studies on hyper-parameters of MCMC-BO. The results are the means and standard deviations of the five replicate independent experiments on Ackley-20d and Rastrigin-20d when the transition number is 300 (The result of the Rastrigin function is shifted by 180).

3. <https://github.com/facebookresearch/nevergrad>

C.4. Low-dimensional Performance

We implement GP-UCB and GP-TS using Gpytorch.

References

- Maximilian Balandat, Brian Karrer, Daniel R. Jiang, Samuel Daulton, Benjamin Letham, Andrew Gordon Wilson, and Eytan Bakshy. BoTorch: A Framework for Efficient Monte-Carlo Bayesian Optimization. In *Advances in Neural Information Processing Systems 33*, 2020.
- Olivier Chapelle and Lihong Li. An empirical evaluation of thompson sampling. In J. Shawe-Taylor, R. Zemel, P. Bartlett, F. Pereira, and K.Q. Weinberger, editors, *Advances in Neural Information Processing Systems*, volume 24. Curran Associates, Inc., 2011.
- Sayak Ray Chowdhury and Aditya Gopalan. On kernelized multi-armed bandits. In *International Conference on Machine Learning*, pages 844–853. PMLR, 2017.
- David Eriksson and Martin Jankowiak. High-dimensional bayesian optimization with sparse axis-aligned subspaces. In *Uncertainty in Artificial Intelligence*, pages 493–503. PMLR, 2021.
- David Eriksson, Michael Pearce, Jacob Gardner, Ryan D Turner, and Matthias Poloczek. Scalable global optimization via local bayesian optimization. *Advances in neural information processing systems*, 32, 2019.
- Peter I Frazier. A tutorial on bayesian optimization. *arXiv preprint arXiv:1807.02811*, 2018.
- Jacob R Gardner, Geoff Pleiss, David Bindel, Kilian Q Weinberger, and Andrew Gordon Wilson. Gpytorch: Blackbox matrix-matrix gaussian process inference with gpu acceleration. In *Advances in Neural Information Processing Systems*, 2018.
- Nikolaus Hansen, Sibylle D Müller, and Petros Koumoutsakos. Reducing the time complexity of the derandomized evolution strategy with covariance matrix adaptation (cma-es). *Evolutionary computation*, 11(1):1–18, 2003.
- James Hensman, Nicolo Fusi, and Neil D Lawrence. Gaussian processes for big data. *arXiv preprint arXiv:1309.6835*, 2013.
- James Hensman, Alexander G Matthews, Maurizio Filippone, and Zoubin Ghahramani. Mcmc for variationally sparse gaussian processes. *Advances in Neural Information Processing Systems*, 28, 2015.
- José Miguel Hernández-Lobato, James Requeima, Edward O Pyzer-Knapp, and Alán Aspuru-Guzik. Parallel and distributed thompson sampling for large-scale accelerated exploration of chemical space. In *International conference on machine learning*, pages 1470–1479. PMLR, 2017.
- Cheng Li, Sunil Gupta, Santu Rana, Vu Nguyen, Svetha Venkatesh, and Alistair Shilton. High dimensional bayesian optimization using dropout. *arXiv preprint arXiv:1802.05400*, 2018.

- Xiuyuan Lu and Benjamin Van Roy. Ensemble sampling. In I. Guyon, U. Von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc., 2017.
- Horia Mania, Aurelia Guy, and Benjamin Recht. Simple random search provides a competitive approach to reinforcement learning. *arXiv preprint arXiv:1803.07055*, 2018.
- Eric Mazumdar, Aldo Pacchiano, Yian Ma, Michael Jordan, and Peter Bartlett. On approximate thompson sampling with Langevin algorithms. In Hal Daumé III and Aarti Singh, editors, *Proceedings of the 37th International Conference on Machine Learning*, volume 119 of *Proceedings of Machine Learning Research*, pages 6797–6807. PMLR, 13–18 Jul 2020.
- Michael D McKay, Richard J Beckman, and William J Conover. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 42(1):55–61, 2000.
- Kerrie L Mengersen and Richard L Tweedie. Rates of convergence of the hastings and metropolis algorithms. *The annals of Statistics*, 24(1):101–121, 1996.
- Nicholas Metropolis, Arianna W Rosenbluth, Marshall N Rosenbluth, Augusta H Teller, and Edward Teller. Equation of state calculations by fast computing machines. *The journal of chemical physics*, 21(6):1087–1092, 1953.
- Amin Nayebi, Alexander Munteanu, and Matthias Poloczek. A framework for bayesian optimization in embedded subspaces. In *International Conference on Machine Learning*, pages 4752–4761. PMLR, 2019.
- ChangYong Oh, Efstratios Gavves, and Max Welling. Bock: Bayesian optimization with cylindrical kernels. In *International Conference on Machine Learning*, pages 3868–3877. PMLR, 2018.
- Carlos Riquelme, G. Tucker, and Jasper Snoek. Deep bayesian bandits showdown: An empirical comparison of bayesian deep networks for thompson sampling. *ArXiv*, abs/1802.09127, 2018.
- Bobak Shahriari, Kevin Swersky, Ziyu Wang, Ryan P Adams, and Nando De Freitas. Taking the human out of the loop: A review of bayesian optimization. *Proceedings of the IEEE*, 104(1):148–175, 2015.
- Jasper Snoek, Hugo Larochelle, and Ryan P Adams. Practical bayesian optimization of machine learning algorithms. *Advances in neural information processing systems*, 25, 2012.
- Il’ya Meerovich Sobol’. On the distribution of points in a cube and the approximate evaluation of integrals. *Zhurnal Vychislitel’noi Matematiki i Matematicheskoi Fiziki*, 7(4):784–802, 1967.
- Emanuel Todorov, Tom Erez, and Yuval Tassa. Mujoco: A physics engine for model-based control. In *2012 IEEE/RSJ international conference on intelligent robots and systems*, pages 5026–5033. IEEE, 2012.
- Iñigo Urteaga and Chris H Wiggins. Variational inference for the multi-armed contextual bandit. *ArXiv*, abs/1709.03163, 2017.

- Linnan Wang, Rodrigo Fonseca, and Yuandong Tian. Learning search space partition for black-box optimization using monte carlo tree search. *Advances in Neural Information Processing Systems*, 33:19511–19522, 2020.
- Zi Wang, Clement Gehring, Pushmeet Kohli, and Stefanie Jegelka. Batched large-scale bayesian optimization in high-dimensional spaces. In *International Conference on Artificial Intelligence and Statistics*, pages 745–754. PMLR, 2018.
- Pan Xu, Hongkai Zheng, Eric V Mazumdar, Kamyar Azizzadenesheli, and Animashree Anandkumar. Langevin Monte Carlo for contextual bandits. In Kamalika Chaudhuri, Stefanie Jegelka, Le Song, Csaba Szepesvari, Gang Niu, and Sivan Sabato, editors, *Proceedings of the 39th International Conference on Machine Learning*, volume 162 of *Proceedings of Machine Learning Research*, pages 24830–24850. PMLR, 17–23 Jul 2022.